

Application No. 09/552,312

Attorney Docket No.: 1999-0097

REMARKS

Reconsideration and allowance are requested. Claims 1 - 17 are pending. Claims 1 and 2 are amended.

Rejection of Claim 1 Under Section 102

The Examiner rejects claim 1 under 35 U.S.C. § 102 in view of U.S. Patent No. 5,703,655 to Corey et al. ("Corey et al."). Applicants respectfully submit that claim 1 as amended is not anticipated by Corey et al. In addition to the earlier arguments presented explaining that Corey et al. do not teach each limitation of claim 1, Applicants submit the following to explain further why amended claim 1 is not taught by the cited prior art reference.

Claim 1 recites a method of indexing media that includes the step of indexing a media collection to create an indexed library based on a content of the media collection. Further, the step of indexing the multimedia content comprises analyzing the media collection to determine whether speech recognition data or closed captioning data may be used to index the media collection. This step is clearly not taught by Corey et al. inasmuch as the reference exclusively focuses on using closed caption data extracted from video content to obtain text for indexing. There is no mention of automatic speech recognition by Corey et al. and further no mention or suggestion of performing an analysis of media content to determine whether speech recognition data or closed captioning data may be used to index the collection. Corey et al. simply assume that closed caption data is carried within the signal. See col. 4, lines 4 - 12. Therefore, Applicants respectfully submit that claim 1 is not taught or suggested by Corey et al. and thus in condition for allowance.

Application No. 09/552,312

Attorney Docket No.: 1999-0097

Rejection of Claims 8 - 13 Under Section 102

The Examiner rejects claims 8 - 13 under 35 U.S.C. § 102 in view of U.S. Patent No. 6,336,093 to Fasciano ("Fasciano"). Applicants traverse this rejection and submit that Fasciano does not anticipate these claims because each limitation is not taught by Fasciano.

We first turn to claim 8. This claim recites a method of indexing media for browsing where a media collection is indexed according to speaker voice characteristics. The next step requires receiving a search query from a user to locate a media segment from the indexed media collection. The Examiner argues that FIG. 7, col. 6, line 63 - col. 7, lines 17 and 40 - 53 match the step of indexing a media collection according to detection of speaker voice characteristics. However, Applicants submit that these portions of Fasciano do not teach this limitation. Fasciano teaches in FIG. 7 and the portions cited notably do not teach indexing a media *collection*. The module in FIG. 7 is disclosed as receiving an audio signal with timecodes and a binary wave form is output. The output signal may be used by an editor to identify splice points in an audio track. The same information may be used to identify points in a video that is being captured. The video program is then divided into segments or clips as it is being processed. This is all the context of video editing. Col. 7, lines 40 - 54 does not disclose indexing a media collection but discloses that satellite feeds of news programs may be captured and automatically directed to a journalists' desk for viewing.

Notably, taking into account the Fasciano context of video editing, the disclosure cited by the Examiner merely relates to using audio for dividing a single video program into segments for the purpose of editing and not browsing. Claim 8 recites a method of indexing media for browsing including the step of indexing a media collection. This recited feature is not taught by Fasciano.

Furthermore, the second step of receiving a search query from a user to locate a media segment from the indexed media collection is not taught by Fasciano. The Examiner equates col. 3, lines 11 - 44 and FIG. 1 with this limitation. However, Fasciano in fact teaches a

Application No. 09/552,312

Attorney Docket No.: 1999-0097

different concept where in the process of video editing, the speech and sound effects are used with a script to match video segments with portions of the script. While editing, a user may provide input to select a portion of the script and thus a video clip portion is selected. The editor first selects a script range and then the editing system displays the script portion and the frames of possible clips for the script as shown at 50. Applicants submit that Fasciano is dealing with the editing of a single video program and thus does not disclose receiving a search query from a user to locate a media segment from an indexed media collection as recited in claim 8. Further, claim 8 clearly recites a method of indexing media for *browsing*, which is within a different context of media editing as is taught by Fasciano. Therefore, claim 8 is in condition for allowance.

Claims 9 - 13 each depend from claim 8 and recite further limitations therefrom. Therefore, since claim 8 is patentable, Applicants submit that these dependent claims are patentable as well.

Rejection of Claims 2 - 4 and 6 Under Section 103

The Examiner rejects claims 2 - 4 and 6 under Section 103 in view of the combination of Corey et al. and Fasciano. Applicants have amended claim 2 into independent form incorporating the original parent claim limitations. Notably, Applicants note that this does not change the scope of original claim 2 and this amendment is not made for patentability since it is of the exact same scope of the original claim 2.

To establish a *prima facie* case of obviousness, the Examiner must meet three criteria. First, there must be some motivation or suggestion, either in the references themselves, or in the knowledge generally available to one of ordinary skill in the art, to combine the references. Second, there must be a reasonable expectation of success, and finally, the prior art references must teach or suggest all the claim limitations. The Examiner bears the initial burden of providing some suggestion of the desirability of doing what the inventor has done.

Application No. 09/552,312

Attorney Docket No.: 1999-0097

"To support the conclusion that the claimed invention is directed to obvious subject matter, either the references must expressly or impliedly suggest the claimed invention or the examiner must present a convincing line of reasoning as to why the artisan would have found the claimed invention to have been obvious in light of the teachings of the references." MPEP 2142.

If a proposed modification would render the prior art invention being modified unsatisfactory for its intended purposes, then there is no suggestion or motivation to make the proposed modification. *In re Gordon*, 733 F.2d 900, 221 USPQ 1125 (Fed. Cir. 1984). Further, if the proposed modification of the prior art would change the principle operation of the prior art invention being modified, then the teaching of the reference is not sufficient to render the claims *prima facie* obvious. *In re Ratti*, 270 F.2d 810, 123 USPQ 349 (CCPA 1959). The principles outlined in both these cases are applicable here.

Applicants agree with the Examiner that Corey et al. fail to teach extracting speech recognition data when closed captioning is unavailable. However, Applicants traverse the Examiner's conclusion that it would be obvious to incorporate the teachings of Fasciano into Corey et al. to provide speech recognition to extract speech patterns. As we have noted above, Corey et al. assumes that closed caption data is available in the modulated data in the video baseband signal. The entire focus of Corey et al. (see Title, Abstract) is to utilize this closed captioning data for indexing purposes. Corey et al. require a closed caption decoder 52 that receives from the baseband video signal 32 the modulated closed captioning data.

In contrast to Corey et al.'s focus on closed captioning text for indexing, Fasciano teaches a method of using speech recognition to capture, author and playback audio and video. The purpose of Fasciano is broader than Corey et al. in that the speech recognition is applied to a number of purposes, such as to enhance speech processing, controlling video capture during editing, and to facilitate clips or splice points for editing. From the title and Abstract of Fasciano, there is no indicating that the concepts disclosed therein would be

Application No. 09/552,312

Attorney Docket No.: 1999-0097

applicable to indexing and retrieval. For example, in col. 2, lines 43 - 51 of Fasciano, he introduces the concept of speech recognition for logging audio material. The context he mentions is logging live audio input from a person providing commentary or for a live event being recorded. Since no closed captioning would exist for such a live speech, Fasciano must rely on speech recognition to obtain text of that live speech.

Applicants respectfully submit that Corey et al. and Fasciano take mutually exclusive approaches to indexing video content. Since Corey et al. rely entirely on the availability of closed captioning data, there is simply no need to incorporate speech recognition capability into their disclosure. Fasciano teaches an exclusive speech/sound recognition method primarily for the purpose of video editing with scant mention of logging and storage of multimedia content based on speech recognition.

If one were to blend the speech recognition concept of Fasciano into the closed captioning invention of Corey et al., the fundamental operating principles of Corey et al. would have to be altered. There would simply be no need for any of the closed captioning apparatus that embodies the core invention of Corey et al. Similarly, if the teachings of Corey et al. were incorporated into Fasciano, then the need for speech recognition would be eliminated since there would be text available for all programming.

Furthermore, there is no suggestion or mention in either reference of expanding into the other's subject matter. Therefore, Applicants respectfully submit that there cannot be any motivation or suggestion either within the references themselves or within the knowledge of one skilled in the art to combine there references. Blending these references would require a modification of the fundamental principles of operation of either Corey et al. or Fasciano. Accordingly, Applicants submit that claims 2 - 4 and 6 are patentable and in condition for allowance.

Application No. 09/552,312

Attorney Docket No.: 1999-0097

Rejection of Claims 5 and 7 Under Section 103

The Examiner rejects claims 5 and 7 under 35 U.S.C. § 103 in view of Corey et al., Fasciano and further in view of U.S. Patent No. 5,983,176 to Hoffert et al. ("Hoffert et al."). Applicants incorporate the above arguments regarding the lack of motivation or suggestion to combine Corey et al. and Fasciano. Inasmuch as the two primary reference that must be combined to support this rejection cannot be legally combined, Applicants submit that claims 5 and 7 are patentable and in condition for allowance for the same reasons.

Rejection of Claims 14 - 17 Under Section 103

The Examiner rejects claims 14 - 17 under 35 U.S.C. § 103 in view of Fasciano and further in view of U.S. Patent No. 6,253,061 to Halverson et al. ("Halverson et al."). Applicants respectfully submit that the parent claim 8 is allowable therefore, dependent claims 14 - 17 are also allowable.

Further, Applicants submit that there is no suggestion or motivation to combine Halverson et al. with Fasciano. Halverson et al. disclose a system for agent-based navigation in a speech-based data navigation system. Their speech-based invention provides for navigating electronic data by using spoken language. See Abstract. Their invention is a front-end spoken language interface to a back-end database. As front-end solution, they disclose interpreting speech, and generating an input template associated with a navigation query into the database. There are several reasons by Halverson et al. should not be combined with Fasciano.

The first reason these references should not be combined is that they are mutually exclusive solutions to different problems. As mentioned above, Fasciano provides a solution to video editors regarding how to capture, author and playback audio and video. The only selection of data occurring in Fasciano occurs when the editor selects portions of a text script

Application No. 09/552,312

Attorney Docket No.: 1999-0097

which automatically matches video segments that are presented to the editor. The text script is generated via automatic speech recognition (ASR) in Fasciano. Halverson et al.'s disclosure relates to a front-end navigation ASR component. Cols. 7 and 8 cited by the Examiner explains step 404 of FIG. 4. As seen in FIG. 4, this step merely interprets a spoken request and provides a preprocessing to the steps of constructing a navigation query (406) and navigating the data source (408). As a front-end navigation solution, Halverson et al. focus on how to receive speech and generate a navigational query to any number of databases. At its core, the invention of Halverson et al. is an ASR database searching method. Inasmuch as Fasciano already internally includes an ASR process to generate the text script for use by the video editor, Applicants respectfully submit that there would be no reason to look outside of Fasciano for other ASR solutions.

Another reason to avoid blending Fasciano with Halverson et al. is that Fasciano, since he focuses on video editing, does not mention back-end databases to be navigated. The "user" in Fasciano is a video editor that selects video segments for generating an audio/video presentation. The way the video segments are identified is through user input according to the text script associated with video clips. The only "database" searched by the editor in Fasciano is the text script of the video program. Therefore, the complexity of the blending these references would be a technological waste. For example, Halverson et al.'s approach requires a complicated process of interpreting a user request, identifying a data source, constructing a navigation query and navigating a data source to obtain a result and transmit the result to the client. This is explained at col. 7, line 10 - col. 12, line 55, all associated with FIG. 4. Such a complicated front-end speech navigation process would simply be unnecessary in Fasciano. The video editor in Fasciano is not navigating large databases of video content, surfing the internet, or facing the "dizzying range of database content choices." Halverson et al., Col. 1, lines 48 - 49. The only searching in Fasciano is shown in FIG. 1 where the editor provides user search input of a selected script range of a single video

Application No. 09/552,312

Attorney Docket No.: 1999-0097

program with the result being a range of video clips. There is certainly no "dizzying range" of content to navigate in Fasciano.

In sum, Applicants submit that there is no suggestion or motivation to combine Fasciano with Halverson et al. for at least these reasons: (1) Fasciano already discloses ASR but fails to mention it as being used by the video editor to search the program script; and (2) blending Fasciano with Halverson et al. would be a technological waste since Fasciano does not disclose navigating multiple databases. Therefore, Applicants submit that claims 14 - 17 are patentable and in condition for allowance.

Rejection of Claims 3 - 7 Under Section 103

The Examiner rejects claims 3 - 7 Under Section 103 as being unpatentable over Corey et al. in view of U.S. Patent No. 6,567,980 to Jain et al. ("Jain et al."). Claims 3 - 7 each depend from claim 2, and recite further limitations therefrom. As mentioned above, Applicants assert that since Corey et al. fail to discuss or disclose any speech recognition data from media content, they cannot anticipate claim 2 that includes limitations of extracting speech recognition data as searchable text data. Accordingly, Applicants submit that claims 3 - 7 are patentable over the prior art of record inasmuch as the primary reference of Corey et al. simply does not disclose each limitation of parent claim 2. Therefore, Corey et al. in view of Jain et al. in combination cannot disclose each limitation of claims 3 - 7.

Therefore, claims 3 - 7 are patentable and in condition for allowance.

Applicants have added new claims 8 - 17. Applicants submit that these claims are patentable and in condition for allowance.

Application No. 09/552,312

Attorney Docket No.: 1999-0097

CONCLUSION

Having addressed the rejection of all the claims, Applicants respectfully submit that the subject application is in condition for allowance and a Notice to that effect is earnestly solicited.

Respectfully submitted,

Date: *July 16, 2004*

Correspondence Address:
Samuel H. Dworetsky
AT&T Corp.
Room 2A-207
One AT&T Way
Bedminster, NJ 07921

By: *Thomas M. Isaacson*

Thomas M. Isaacson
Attorney for Applicants
Reg. No. 44,166
Phone: 410-414-3056
Fax No.: 410-510-1433